ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 33, Number 4, Winter 2003

INVERTIBILITY CONDITIONS FOR BLOCK MATRICES AND ESTIMATES FOR NORMS OF INVERSE MATRICES

M.I. GIL'

ABSTRACT. A nonsingularity criterion for block matrices is derived. It improves the well-known results in the case of matrices which are "close" to block triangular ones. Moreover, an estimate for the norm of the inverse matrices is derived.

1. Introduction and statement of the main result. Although excellent computer software are now available for eigenvalue computation, new results on invertibility and spectrum inclusion regions for finite matrices are still important, since computers are not very useful, in particular, for analysis of matrices dependent on parameters. But such matrices play an essential role in various applications, for example, in stability and boundedness of coupled systems of partial differential equations, cf. [**9**, Section 14]. In addition, invertibility conditions for finite matrices allow us to derive invertibility conditions for linear operators and, in particular, infinite matrices. Because of this, the problem of finding invertibility conditions and spectrum inclusion regions for finite matrices continues to attract the attention of many specialists, cf. [**1–3, 8, 10–12, 14**] and references given therein.

Many books and papers are devoted to the invertibility of block matrices [**4, 5, 6, 13**], etc. In these works mainly, the Hadamard theorem is generalized to block matrices. Note that the generalized Hadamard theorem does not assert that a block triangular matrix with nonsingular diagonal blocks is invertible. But it is not hard to check that such a matrix is always invertible. In the present paper, we propose invertibility conditions which improve well-known results for matrices that are "close" to block triangular matrices. Moreover, we derive an estimate for the norm of the inverse matrices.

AMS *Mathematics Subject Classification.* 15A9, 15A18.
Key words and phrases. Block matrices, nonsingularity.
This research was supported by the Israel Ministry of Science and Technology
Received by the editors on August 10, 2000, and in revised form on August 15, 2001.

Let $n \times n$-matrix $A = (a_{jk})_{j,k=1}^n$ be partitioned in the following manner:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ . & \cdots & . & . \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{pmatrix}, \tag{1.1}$$

where $m < n$, A_{jk} are matrices. Below $I = I_n$ is the unit operator in $\mathbf{C}^n$. As usual, $\|A\|_\infty$ is the norm defined by

$$\|A\|_\infty = \max_{j=1,\dots,n} \sum_{k=1}^n |a_{jk}|.$$

Let the diagonal blocks A_{kk} be invertible. Denote

$$v_k^{up} = \max_{j=1,2,\dots,k-1} \|A_{jk}A_{kk}^{-1}\|_\infty, \quad k = 2,\dots,m$$
$$v_k^{low} = \max_{j=k+1,\dots,m} \|A_{jk}A_{kk}^{-1}\|_\infty, \quad k = 1,\dots,m-1.$$

The aim of the present paper is to prove the following

Theorem 1.1. *Let the diagonal blocks A_{kk}, $k = 1,\dots,m$, be invertible. In addition, with the notations*

$$M_{up} \equiv \prod_{2\le k\le m} (1 + v_k^{up}), \quad M_{low} \equiv \prod_{1\le k\le m-1} (1 + v_k^{low}),$$

let the condition

$$M_{low}M_{up} < M_{low} + M_{up} \tag{1.2}$$

hold. Then matrix A defined by (1.1) *is invertible. Moreover,*

$$\|A^{-1}\|_\infty \le \frac{\max_k \|A_{kk}^{-1}\|_\infty M_{low}M_{up}}{M_{low} + M_{up} - M_{low}M_{up}}. \tag{1.3}$$

The proof of this theorem is divided into a series of lemmas, which are presented in the next sections.

Theorem 1.1 is exact in the following sense. Let the matrix A in (1.1) be upper block triangular and let A_{kk} be invertible. Then $M_{low} = 1$ and condition (1.2) takes the form $M_{up} < 1 + M_{up}$. Thus, due to Theorem 1.1, A is invertible. We have the same result if the matrix in (1.1) is upper block triangular.

Recall that the generalized Hadamard theorem, cf. [**6**, Section 14.3, Theorem 3] gives the following invertibility conditions for the matrix in (1.1) (in the terms of norm $\|\cdot\|_\infty$):

$$\|A_{jj}^{-1}\|_\infty^{-1} > \sum_{\substack{k=1\\k\neq j}}^{m} \|A_{jk}\|_\infty, \quad j = 1, \dots, m. \tag{1.4}$$

Let us compare Theorem 1.1 with condition (1.4). Rewrite (1.4) as

$$\theta_H \equiv \max_{j=1,\dots,m} \|A_{jj}^{-1}\|_\infty \sum_{\substack{k=1\\k\neq j}}^{m} \|A_{jk}\|_\infty < 1$$

and note that (1.2) can be rewritten as

$$(M_{low} - 1)(M_{up} - 1) < 1. \tag{1.5}$$

Thus, Theorem 1.1 improves condition (1.4) if

$$(M_{low} - 1)(M_{up} - 1) < \theta_H. \tag{1.6}$$

Consider now a block matrix with $m = 2$:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{1.7}$$

Then $v_2^{up} = \|A_{12}A_{22}^{-1}\|_\infty$, $v_1^{low} = \|A_{21}A_{11}^{-1}\|_\infty$, $M_{up} = 1 + v_2^{up}$ and $M_{low} = 1 + v_1^{low}$. Assume that

$$(1 + v_2^{up})(1 + v_1^{low}) < 2 + v_2^{up} + v_1^{low},$$

or equivalently,

$$\|A_{12}A_{22}^{-1}\|_\infty \|A_{21}A_{11}^{-1}\|_\infty < 1. \tag{1.8}$$

Then, due to Theorem 1.1, the matrix in (1.7) is invertible. Moreover,

$$\|A^{-1}\|_\infty \le \frac{\max_{k=1,2} \|A_{kk}^{-1}\|_\infty M_{low} M_{up}}{M_{low} + M_{up} - M_{low} M_{up}}.$$

At the same time, condition (1.4) takes the form

$$\|A_{11}^{-1}\|_\infty \|A_{12}\|_\infty < 1, \quad \|A_{22}^{-1}\|_\infty \|A_{21}\|_\infty < 1.$$

Certainly (1.8) is sharper than the latter inequalities.

Assume now that m is even: $m = 2m_0$ with a natural m_0, and A_{jk} are 2×2 matrices:

$$A_{jk} = \begin{pmatrix} a_{2j-1,2k-1} & a_{2j-1,2k} \\ -a_{2j-1,2k} & a_{2j,2k} \end{pmatrix}$$

with $j, k = 1, \cdots, m_0$. Take into account that

$$A_{kk}^{-1} = \begin{pmatrix} a_{2k,2k} & -a_{2k-1,2k} \\ -a_{2k,2k-1} & a_{2k-1,2k-1} \end{pmatrix}$$

with

$$d_k = a_{2k-1,2k-1} a_{2k,2k} - a_{2k,2k-1} a_{2k-1,2k}.$$

Thus, the quantities v_k^{up}, v_k^{low} are simple to calculate. Now relation (1.2) yields the invertibility and (1.3) gives the estimate for the inverse matrix.

In Section 4 below we show that the matrix in (1.7) is invertible provided

$$\|A_{12} A_{22}^{-1} A_{21} A_{11}^{-1}\| < 1 \tag{1.9}$$

with an arbitrary matrix norm $\|\cdot\|$ in $\mathbf{C}^n$. This result slightly improves condition (1.8).

2. π-triangular matrices.

Let $B(\mathbf{C}^n)$ be the set of all linear operators in $\mathbf{C}^n$. In what follows $\pi = \{P_k, k = 0, \dots, m \le n\}$ is a *chain of orthogonal projectors P_k in $\mathbf{C}^n$, such that*

$$0 = P_0 \subset P_1 \subset \cdots \subset P_m = I - n.$$

The relation $P_{k-1} \subset P_k$ means that $P_{k-1}\mathbf{C}^n \subset P_k\mathbf{C}^n$, $k = 1, \dots, m$.

Let operators $A, D, V \in B(C^n)$ satisfy the relations

$$AP_k = P_kAP_k, \quad k = 1, \dots, m, \tag{2.1}$$

$$DP_k = P_kD, \quad k = 1, \dots, m, \tag{2.2}$$

$$VP_k = P_{k-1}VP_k, \quad k = 2, \dots, m; \quad VP_1 = 0. \tag{2.3}$$

Then A, D and V will be called a π-*triangular operator*, a π-*diagonal one* and π-*nilpotent* one, respectively.

Since

$$\begin{aligned} V^m &= V^mP_m = V^{m-1}P_{m-1}V = V^{m-2}P_{m-2}VP_{m-1}V = \dots \\ &= VP_1 \dots VP_{m-2}VP_{m-1}V, \end{aligned}$$

we have

$$V^m = 0. \tag{2.4}$$

That is, every π-nilpotent operator is a nilpotent operator. Denote

$$\Delta P_k = P_k - P_{k-1}; \quad V_k = V\Delta P_k, \quad k = 1, \dots, m.$$

Lemma 2.1. *Let A be π-triangular. Then*

$$A = D + V, \tag{2.5}$$

here V is a π-nilpotent operator and D is a π-diagonal one.

Proof. Clearly,

$$A = \sum_{j=1}^{m} \Delta P_j A \sum_{k=1}^{m} \Delta P_k = A = \sum_{k=1}^{m}\sum_{j=1}^{k} \Delta P_j A \Delta P_k = \sum_{k=1}^{m} P_k A \Delta P_k.$$

Hence, (2.5) is valid with

$$D = \sum_{k=1}^{m} \Delta P_k A \Delta P_k, \tag{2.6}$$

and

$$V = \sum_{k=2}^{m} P_{k-1} A \Delta P_k = \sum_{k=2}^{m} V_k. \quad \square \tag{2.7}$$

Definition 2.2. Let $A \in B(\mathbf{C}^n)$ be a π-triangular operator and suppose (2.5) holds. Then the π-diagonal operator D and the π-nilpotent operator V will be called the diagonal part of A and nilpotent part of A, respectively.

Lemma 2.3. *Let π be a chain of orthogonal projectors in $\mathbf{C}^n$. If $\tilde{V}$ is a π-nilpotent operator, A is a π-triangular one, then both operators $A\tilde{V}$ and $\tilde{V}A$ are π-nilpotent ones.*

Proof. By (2.1) and (2.3), we get

$$\tilde{V}AP_k = \tilde{V}P_kAP_k = P_{k-1}\tilde{V}P_kAP_k = P_{k-1}\tilde{V}AP_k, \quad k = 1, \dots, m.$$

That is, $\tilde{V}A$ indeed is a π-nilpotent operator. Similarly we can prove that $A\tilde{V}$ is a π-nilpotent operator. $\square$

Lemma 2.4. *Let $A \in B(\mathbf{C}^n)$ be a π-triangular operator and D be its diagonal part. Then the spectrum $\sigma(A)$ of A coincides with the spectrum of $\sigma(D)$ of D.*

Proof. Due to (2.6),

$$R_\lambda(A) = (A - \lambda I)^{-1} = (D + V - \lambda I)^{-1} = R_\lambda(D)(I + VR_\lambda(D))^{-1}. \tag{2.8}$$

According to Lemma 2.3, $VR_\lambda(D)$ is π-nilpotent if λ is not an eigenvalue of D. Therefore

$$R_\lambda(D)(I + VR_\lambda(D))^{-1} = \sum_{k=0}^{m-1} R_\lambda(D)(-VR_\lambda(D))^k.$$

This relation implies that λ is a regular point of A if it is a regular point of D. That is, $\lambda \notin \sigma(A)$ if $\lambda \notin \sigma(D)$. Besides, if $\lambda \in \sigma(D)$, then $\lambda \in \sigma(A)$. This is precisely the assertion of the lemma. $\square$

3. Multiplicative representation of resolvents of π-triangular operators. For $X_1, X_2, \dots, X_m \in B(\mathbf{C}^n)$ and $j < m$, denote

$$\overrightarrow{\prod_{j\le k\le m}} X_k \equiv X_j X_{j+1} \dots X_m$$

(the arrow over the symbol of the product means that the indexes of the co-factors increase from left to right).

Lemma 3.1. *Let V be a π-nilpotent operator; then*

$$(I-V)^{-1} = \overrightarrow{\prod_{2\le k\le m}} (I + V\Delta P_k). \tag{3.1}$$

Proof. According to (2.4),

$$(I-V)^{-1} = \sum_{k=0}^{m-1} V^k. \tag{3.2}$$

On the other hand,

$$\overrightarrow{\prod_{2\le k\le m}} (I + V\Delta P_k) = I + \sum_{k=2}^{m} V_k + \sum_{s\le k_1<k_2\le m-1} V_{k_1}V_{k_2} + \sum_{2\le k_1<k_3\dots<k_{m-1}\le m} V_{k_1}V_{k_2}\dots V_{k_m}.$$

Here, as above, $V_k = V\Delta P_k$. But

$$\sum_{2\le k_1<k_2\le m} V_{k_1}V_{k_2} = V$$

$$\sum_{2\le k_1<k_2\le m} \Delta P_{k_1} V \Delta P_{k_2} = V \sum_{3\le k_2\le m} P_{k-1} V \Delta P_{k_2} = V^2 \sum_{3\le k_2\le m} \Delta P_{k_2} = V^2.$$

Similarly,

$$\sum_{2\le k_1<k_3\cdots<k_{j-1}\le m} V_{k_1}V_{k_2}\dots V_{k_j} = V^j \text{ for } j<m.$$

Thus, from (3.2), (3.1) follows. This is the desired conclusion. □

Theorem 3.2. *For any π-triangular operator A and a regular $\lambda \in \mathcal{C}$*

$$R_\lambda(A) = (D-\lambda I)^{-1} \overrightarrow{\prod_{2\le k\le m}} (I - V_k(A_{kk}-\lambda \Delta P_k)^{-1}\Delta P_k),$$

where D and V are the π-diagonal and π-nilpotent parts of A, respectively.

Proof. Due to Lemma 2.3, $VR_\lambda(D)$ is π-nilpotent. Now Lemma 3.1 gives

$$(I+VR_\lambda(D))^{-1} = \overrightarrow{\prod_{2\le k\le m}} (I - VR_\lambda(D)\Delta P_k).$$

But

$$R_\lambda(D)\Delta P_k = \Delta P_k(A_{kk}-\lambda\Delta P_k)^{-1}.$$

This and (2.8) prove the result. □

Note that Theorem 3.2 is a particular generalization of Theorem 1.6.1 from [**7**].

4. Invertibility with respect to a chain of properties. Let again $\pi = \{P_k, k=0,\dots,m\}$ be a chain of orthogonal projectors P_k. Denote by $\tilde\pi$ the chain of the complementary projectors $\tilde P_k = I_n - P_{m-k} : \tilde\pi = \{I_n - P_{m-k}, k=0,\dots,m\}$.

In the present section, V is a π-nilpotent operator, D is a π-diagonal one, and W is $\tilde\pi$-nilpotent operator.

Lemma 4.1. *Any operator $A\in B(\mathbf{C}^n)$ admits the representation*

$$A = D + V + W. \tag{4.1}$$

Proof. Clearly,

$$A = \sum_{j=1}^{m} \Delta P_j A \sum_{k=1}^{m} \Delta P_k.$$

Hence (4.1) holds, where D and V are defined by (2.6) and (2.7), and

$$W = \sum_{k=1}^{m} \sum_{j=k+1}^{m} \Delta P_j A \Delta P_k = \sum_{k=2}^{m} \tilde{P}_{k-1} A \Delta \tilde{P}_k$$

with $\Delta \tilde{P}_k = P_k - P_{k-1}$. $\square$

Let $\|\cdot\|$ be an arbitrary norm in $\mathbf{C}^n$.

Lemma 4.2. *Let the* π*-diagonal matrix* D *be invertible. In addition, with the notations*

$$V_A \equiv D^{-1} V, \quad W_A \equiv D^{-1} W,$$

let the condition

$$\theta \equiv \|V_A (I + V_A)^{-1} (I + W_A)^{-1} W_A\| < 1 \tag{4.2}$$

hold. Then the operator A *defined by* (4.1) *is invertible. Moreover,*

$$\|A^{-1}\| \le \|(I + V_A)^{-1} D^{-1}\| \|(I + W_A)^{-1}\| (1 - \theta)^{-1}. \tag{4.3}$$

Proof. Due to Lemma 4.1,

$$A = D + V + W = D(I + V_A + W_A) = D[(I + V_A)(I + W_A) - V_A W_A].$$

Clearly, V_A and W_A are nilpotent matrices and, consequently the matrices $I + V_A$ and $I + W_A$ are invertible. Thus,

$$A = D(I + V_A)(I - B_A)(I + W_A), \tag{4.4}$$

where

$$B_A = (I + V_A)^{-1} V_A W_A (I + W_A)^{-1}.$$

Condition (4.2) yields

$$\|(I - B_A)^{-1}\| \le (1-\theta)^{-1}.$$

Therefore, (4.2) provides the invertibility of A. Moreover, according to (4.4), inequality (4.3) is valid. $\square$

Corollary 4.3. *Under condition* (1.9)*, the matrix in* (1.7) *is invertible.*

Indeed, under the consideration $V_A^2 = 0$, $W_A^2 = 0$. So $V_A(I+V_A)^{-1} = V_A$ and $W_A(I+W_A)^{-1} = W_A$. Hence $\theta = \|V_A W_A\|$. Now Lemma 4.2 yields the required result.

Furthermore, Lemmas 3.1 and 2.3 yield the following.

Lemma 4.4. *The inequalities*

$$\|(I+V_A)^{-1}\| \le M(V_A) \equiv \prod_{2\le k\le m} (1 + \|VD^{-1}\Delta P_k\|)$$

and

$$\|(I+W_A)^{-1}\| \le M(W_A) \equiv \prod_{1\le k\le m-1} (1 + \|WD^{-1}\Delta P_k\|)$$

are valid.

5. Proof of Theorem 1.1. In the present section D, V and W are the diagonal, upper diagonal and lower diagonal parts of the matrix in (1.1), respectively, that is,

(5.1)

$$D = \begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ . & \cdots & . & . \\ 0 & 0 & \cdots & A_{mm} \end{pmatrix}, \quad V = \begin{pmatrix} 0 & A_{12} & \cdots & A_{1m} \\ 0 & 0 & \cdots & A_{2m} \\ . & \cdots & . & . \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

and

$$W = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ A_{21} & 0 & \cdots & 0 \\ . & \cdots & . & . \\ A_{m1} & A_{m2} & \cdots & 0 \end{pmatrix}. \tag{5.2}$$

Recall that $V_A \equiv D^{-1}V, \quad W_A \equiv D^{-1}W$.

Lemma 5.1. *Let D, V and W be as in* (5.1) *and* (5.2)*, and let D be invertible. Then the inequalities*

$$\|(I + V_A)^{-1}\|_\infty \leq M_{up} \tag{5.3}$$

and

$$\|(I + W_A)^{-1}\|_\infty \leq M_{low} \tag{5.4}$$

are valid.

Proof. Let $\hat{\pi} = \{\hat{P}_k, k = 1, \dots, m\}$ be the chain of the projectors onto the standard basis. That is, for an $h = (h_k) \in \mathbf{C}^n$,

$$\hat{P}_k h = (h_1, \dots, h_{\nu_k}, \dots, 0),$$

where $\nu_k = \dim \hat{P}_k$. Then according (5.1), D and V are $\hat{\pi}$-diagonal and $\hat{\pi}$-nilpotent operators, respectively. Moreover,

$$VD^{-1}\Delta\hat{P}_k = V_k A_{kk}^{-1}\Delta\hat{P}_k = \sum_{j=1}^{k-1} A_{jk}A_{kk}^{-1}\Delta\hat{P}_k.$$

But, clearly,

$$\left\| \sum_{j=1}^{k-1} A_{jk}A_{kk}^{-1}\Delta\hat{P}_k \right\|_\infty = \max_j \|A_{jk}A_{kk}^{-1}\|_\infty = v_k^{up}.$$

Therefore, inequality (5.3) is due to the previous lemma. Inequality (5.4) can be proved similarly. □

Lemma 5.2. *Let D, V and W be as in* (5.1) *and* (5.2) *and D invertible. Then the inequalities*

$$\|V_A(I + V_A)^{-1}\|_\infty \leq M_{up} - 1 \tag{5.5}$$

and

$$\|W_A(I + W_A)^{-1}\|_\infty \leq M_{low} - 1 \tag{5.6}$$

are valid.

Proof. Let $B = (b_{jk})_{k=1}^{n}$ be a positive matrix with the property

$$Bh \geq h \tag{5.7}$$

for any nonnegative $h \in \mathbf{C}^n$. Then

$$\|B - I\|_\infty = \max_{j=1,\dots,n} \left[\sum_{k=1}^{n} b_{jk} - \delta_{jk} \right] = \|B\|_\infty - 1. \tag{5.8}$$

Here δ_{jk} is the Kronecker symbol. Furthermore, since V_A is nilpotent,

$$\|V_A(I + V_A)^{-1}\|_\infty \leq \left\| \sum_{k=1}^{n-1} |V_A|^k \right\|_\infty = \|(I - |V_A|)^{-1} - I\|_\infty,$$

where $|V_A|$ is the matrix whose entries are the absolute values of the entries of V_A. Moreover,

$$\sum_{k=0}^{n-1} |V_A|^k h \geq h$$

for any nonnegative $h \in \mathbf{C}^n$. So, according to (5.7) and (5.8)

$$\|V_A(I + V_A)^{-1}\|_\infty \leq \|(I - |V_A|)^{-1})^{-1}\|_\infty - 1.$$

Since $\||V_A|\Delta \hat{P}_k\|_\infty = \|V_A \Delta \hat{P}_k\|_\infty$, inequality (5.3) with $-|V_A|$ instead of V_A yields inequality (5.5). Inequality (5.6) can be proved similarly. $\square$

The assertion of Theorem 1.1 follows from Lemmas 4.2, 5.1 and 5.2. $\square$

REFERENCES

1. R.A. Brualdi, *Matrices, eigenvalues and directed graphs*, Linear and Multilinear Algebra **11** (1982), 143–165.

2. F.O. Farid, *Criteria for invertibility of diagonally dominant matrices*, Linear Algebra Appl. **215** (1995), 63–93.

3. ———, *Topics on a generalization of Gershgorin's theorem*, Linear Algebra Appl. **268** (1998), 91–116.

4. D.G. Feingold and R.S. Varga, *Block diagonally dominant matrices and generalization of the Gershgorin circle theorem*, Pacific J. Math. **12** (1962), 1241–1250.

5. M. Fiedler, *Some estimates of spectra of matrices*, Symposium of the Numerical Treatment of Ordinary Differential Equations, Integral and Integro-Differential Equations (Rome, 1960), Birkhaüser Verlag, 1960, 33–36.

6. F.R. Gantmacher, *Theory of matrices*, Nauka, Moscow, 1967 (in Russian).

7. M.I. Gil', *Norm estimations for operator-valued functions and applications*, Marcel Dekker, Inc., New York, 1995.

8. ———, *A nonsingularity criterion for matrices*, Linear Algebra Appl. **253** (1997), 79–87.

9. ———, *Stability of finite and infinite dimensional systems*, Kluwer Acad. Publ., Dordrecht, 1995.

10. R.A. Horn and C.R. Johnson, *Topics in matrix analysis*, Cambridge Univ. Press, Cambridge, 1991.

11. Bishan Li and M.J. Tsatsomeros, *Doubly diagonally dominant matrices*, Linear Algebra Appl. **261** (1997), 221–235.

12. M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, 1964.

13. A.M. Ostrowski, *On some metrical properties of operator matrices and matrices partitioned into blocks.*, J. Math. Ann. Appl. **2** (1961), 161–209.

14. B.S. Tam, S. Yang and X. Zhang, *Invertibility of irreducible matrices*, Linear Algebra Appl. **259** (1997), 39–70.

Department of Mathematics, Ben Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel
E-mail address: gilmi@cs.bgu.ac.il